

1        19. A system for generalizing a set of atomics and/or groups in a hierarchical document  
2        structure, the system comprising:

3            means for identifying an anchor node, the anchor node being a context node of a template for  
4        a particular node of content;

5            means for identifying an anchor node parent with sibling case where particular nodes of  
6        content share the same anchor node and the path expressions for each particular node of content are  
7        the same as the anchor node;

8            means for determining the anchors if the anchor node parent with sibling case is identified;

9            means for combining the location expressions of the identified nodes of content into a single  
10      location expression for a generalized anchor node;

11            means for determining if the generalized anchor node is a sibling; and

12            means for generating a generalized expression corresponding to the generalized anchor node  
13      that locates the content in the particular nodes of content identified.

1        20. The system of Claim 19 further comprising means for reanchoring the particular  
2        nodes of content to a reanchor node if the generalized anchor node is a sibling node and means for  
3        determining if the reanchor node is tangled such that the location expression to a piece of content  
4        matches more than one piece of content.

1        21. The system of Claim 19 further comprising means for identifying the lowest node in  
2        the hierarchical document structure that has not been generalized and means for generalizing the  
3        lowest node before generalizing the nodes that are higher in the hierarchical document structure.

1        22. The system of Claim 19, wherein the location expression combining means further  
2        comprises means for identifying a location expression for each particular node of content, means for  
3        determining if there are other nodes of content and means for generating a replacement anchor node  
4        if there are no other nodes of content.

1        23. The system of Claim 22, wherein the location expression combining means further  
2        comprises means for determining if the location expression for the other nodes of content have been  
3        generalized, means for generalizing the location expressions of the other nodes of content if they  
4        have not been previously generalized and means for identifying the previously generalized location  
5        expressions.

1        24. The system of Claim 23, wherein the location expression combining means further  
2 comprises means for determining if the code associated with the location expression are consistent  
3 with each other, means for generalizing each element of a location expression if the code is not  
4 consistent and means for generalizing the common elements in the path if the code is consistent.

1        25. The system of Claim 20, wherein the means for determining a tangled node further  
2 comprises means for determining the anchor nodes in the hierarchical document structure and means  
3 for generating replacement nodes for location expressions having the same number of elements if  
4 there are no more anchor nodes.

1        26. The system of Claim 25, wherein the means for determining a tangled node further  
2 comprises means for determining the number of elements in each location expression and means for  
3 indexing each location expression according to location, anchor number and element number.

1        27. A method for generalizing a set of atomics and/or groups in a hierarchical document  
2 structure, the method comprising:

3              identifying an anchor node, the anchor node being a context node of a template for a  
4 particular node of content;

5              identifying an anchor node parent with sibling case where particular nodes of content share  
6 the same anchor node and the path expressions for each particular node of content are the same as  
7 the anchor node;

8              determining the anchors if the anchor node parent with sibling case is identified;

9              combining the location expressions of the identified nodes of content into a single location  
10 expression for a generalized anchor node;

11              determining if the generalized anchor node is a sibling; and

12              generating a generalized expression corresponding to the generalized anchor node that locates  
13 the content in the particular nodes of content identified.

1        28. The method of Claim 27 further comprising reanchoring the particular nodes of  
2 content to a reanchor node if the generalized anchor node is a sibling node and determining if the  
3 reanchor node is tangled such that the location expression to a piece of content matches more than  
4 one piece of content.

1           29. The method of Claim 27 further comprising identifying the lowest node in the  
 2 hierarchical document structure that has not been generalized and generalizing the lowest node  
 3 before generalizing the nodes that are higher in the hierarchical document structure.

1           30. The method of Claim 27, wherein the location expression combining further  
 2 comprises identifying a location expression for each particular node of content, determining if there  
 3 are other nodes of content and generating a replacement anchor node if there are no other nodes of  
 4 content.

1           31. The method of Claim 30, wherein the location expression combining further  
 2 comprises determining if the location expression for the other nodes of content have been  
 3 generalized, generalizing the location expressions of the other nodes of content if they have not been  
 4 previously generalized and identifying the previously generalized location expressions.

1           32. The method of Claim 31, wherein the location expression combining further  
 2 comprises determining if the code associated with the location expression are consistent with each  
 3 other, generalizing each element of a location expression if the code is not consistent and  
 4 generalizing the common elements in the path if the code is consistent.

1           33. The method of Claim 28, wherein determining a tangled node further comprises  
 2 determining the anchor nodes in the hierarchical document structure and generating replacement  
 3 nodes for location expressions having the same number of elements if there are no more anchor  
 4 nodes.

1           34. The method of Claim 33, wherein the determining a tangled node further comprises  
 2 determining the number of elements in each location expression and indexing each location  
 3 expression according to location, anchor number and element number.

1           35. A system for generalizing a set of atomics and/or groups in a hierarchical document  
 2 structure, the system comprising:

3           means for identifying an anchor node, the anchor node being a context XHTML node of the  
 4 XSL template for a particular RML node;

5           means for identifying an anchor node parent with sibling delimiters where, each item shares  
 6 the same parent;

7           means for identifying an anchor node sibling where, each individual area of generalized  
 8 structure is not capable of being contained underneath its own unique ancestor node;

TOP SECRET//DTI//REL TO GOVT